



(12)

Veröffentlichung

der internationalen Anmeldung mit der
(87) Veröffentlichungs-Nr.: **WO 2022/190531**
in der deutschen Übersetzung (Art. III § 8 Abs. 2
IntPatÜbkG)
(21) Deutsches Aktenzeichen: **11 2021 007 212.9**
(86) PCT-Aktenzeichen: **PCT/JP2021/047100**
(86) PCT-Anmeldetag: **20.12.2021**
(87) PCT-Veröffentlichungstag: **15.09.2022**
(43) Veröffentlichungstag der PCT Anmeldung
in deutscher Übersetzung: **04.01.2024**

(51) Int Cl.: **G06V 10/22 (2022.01)**
G06T 7/00 (2017.01)

(30) Unionspriorität:
2021-036637 08.03.2021 JP

(72) Erfinder:
Sakata, Shinya, Kyoto, JP

(71) Anmelder:
Omron Corporation, Kyoto, JP

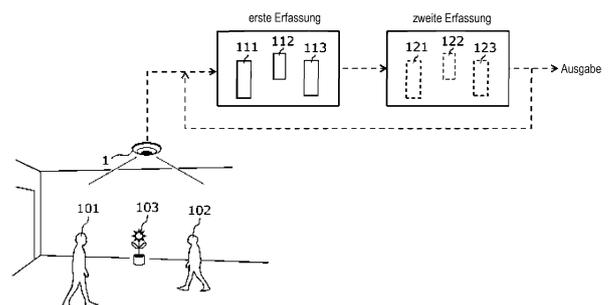
(74) Vertreter:
**isarpatent - Patent- und Rechtsanwälte Barth
Hassa Peckmann und Partner mbB, 80801
München, DE**

Prüfungsantrag gemäß § 44 PatG ist gestellt.

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.

(54) Bezeichnung: **Objekterfassungseinrichtung, Objekterfassungsverfahren und Programm**

(57) Zusammenfassung: Eine Objekterfassungseinrichtung zum Erfassen eines vorbestimmten Objekts aus einem Bild beinhaltet einen ersten Detektor, der einen oder mehrere Kandidatenbereiche einschließlich des vorbestimmten Objekts aus dem Bild erfasst, eine Bestimmungseinheit, die einen Zielbereich aus dem einen oder den mehreren durch den ersten Detektor erfassten Kandidatenbereichen bestimmt, einen zweiten Detektor, der das vorbestimmte Objekt in dem Zielbereich mit einem Erfassungsalgorithmus erfasst, der sich von einem durch den ersten Detektor verwendeten Erfassungsalgorithmus unterscheidet, und einen Speicher, der Erfassungsinformationen speichert, die ein durch den zweiten Detektor für den Zielbereich erhaltenes Erfassungsergebnis angeben. Die Bestimmungseinheit bestimmt den Zielbereich aus dem einen oder den mehreren Kandidatenbereichen basierend auf den Erfassungsinformationen über einen Rahmen, der mindestens ein Rahmen ist, der einem aktuellen Rahmen vorausgeht.



Beschreibung

TECHNISCHES GEBIET

[0001] Die vorliegende Erfindung betrifft eine Technik zum Erfassen eines Objekts.

HINTERGRUND DER ERFINDUNG

[0002] Bekannte Objekterfassungstechniken können zwei Detektoren verwenden, die ein vorhergehender Detektor und ein nachfolgender Detektor sind. Beispielsweise wird die in den Patentliteraturen 1 und 2 beschriebene Objekterfassung mit hoher Genauigkeit unter Verwendung eines vorhergehenden Detektors, der Kandidatenbereiche für Erfassungsziele (z. B. das Gesicht) erfasst, und eines nachfolgenden Detektors, der die Erfassungsziele aus den mehreren Kandidatenbereichen erfasst, durchgeführt.

LITERATURVERZEICHNIS

PATENTLITERATUR

Patentliteratur 1: Japanische Offenlegungsschrift der ungeprüften Patentanmeldung Nr. 2006-293720

Patentliteratur 2: Japanische Offenlegungsschrift der ungeprüften Patentanmeldung Nr. 2019-021001

ZUSAMMENFASSUNG

TECHNISCHES PROBLEM

[0003] Solche bekannten Techniken führen jedoch zwei separate Erfassungsprozesse oder einen zweistufigen Erfassungsprozess durch, die mehr Verarbeitungszeit verwenden. Insbesondere wenn eine stationäre Kamera verwendet wird, erscheinen Objekte mit Ausnahme von Erfassungszielobjekten (z. B. sich bewegende Objekte) auf die gleiche Weise. Somit wird eine einmal durch den vorhergehenden Detektor durchgeführte fehlerhafte Erfassung an der gleichen Stelle wiederholt. Insbesondere führt der nachfolgende Detektor einen Erfassungsprozess für den Bereich durch, für den eine fehlerhafte Erfassung durchgeführt wurde, wodurch die Verarbeitungszeit weiter erhöht wird.

[0004] Ein oder mehrere Aspekte der vorliegenden Erfindung sind auf eine Technik zur Objekterfassung mit hoher Geschwindigkeit und mit hoher Genauigkeit gerichtet.

LÖSUNG DER AUFGABE

[0005] Die Technik gemäß einem oder mehreren Aspekten der vorliegenden Erfindung stellt die nachstehende Struktur bereit.

[0006] Eine Objekterfassungseinrichtung gemäß einem ersten Aspekt der vorliegenden Erfindung erfasst ein vorbestimmtes Objekt aus einem Bild. Die Einrichtung beinhaltet einen ersten Detektor, der einen oder mehrere Kandidatenbereiche einschließlich des vorbestimmten Objekts aus dem Bild erfasst, eine Bestimmungseinheit, die einen Zielbereich aus dem einen oder den mehreren durch den ersten Detektor erfassten Kandidatenbereichen bestimmt, einen zweiten Detektor, der das vorbestimmte Objekt in dem Zielbereich mit einem Erfassungsalgorithmus erfasst, der sich von einem durch den ersten Detektor verwendeten Erfassungsalgorithmus unterscheidet, und einen Speicher, der Erfassungsinformationen speichert, die ein durch den zweiten Detektor für den Zielbereich erhaltenes Erfassungsergebnis angeben. Die Bestimmungseinheit bestimmt den Zielbereich aus dem einen oder den mehreren Kandidatenbereichen basierend auf den Erfassungsinformationen über einen Rahmen, der mindestens ein Rahmen ist, der einem aktuellen Rahmen vorausgeht.

[0007] Beliebige Objekte können Erfassungsziele sein, einschließlich eines menschlichen Körpers, eines Gesichts, eines bestimmten Tiers, eines Automobils und eines bestimmten Produkts. Ein Kandidatenbereich wird mit dem ersten Detektor als wahrscheinlich ein Erfassungszielobjekt enthaltend bestimmt. Dieser Kandidatenbereich wird verwendet, um einen Bereich (Zielbereich) zu bestimmen, für den die Erfassung mit dem zweiten Detektor durchgeführt wird. Obwohl der erste Detektor und der zweite Detektor einen beliebigen Algorithmus verwenden können, kann der durch den zweiten Detektor verwendete Erfassungsalgorithmus eine genauere Erfassung mit einer größeren Rechenmenge als der durch den ersten Detektor verwendete Erfassungsalgorithmus ermöglichen. Die Objekterfassung mit dem zweiten Detektor erzeugt Erfassungsinformationen, einschließlich beispielsweise der Position und der Größe eines Zielbereichs, eines Bilds, das dem Zielbereich entspricht, und einer Bewertung, die die Wahrscheinlichkeit angibt, dass das Erfassungszielobjekt in dem Zielbereich enthalten ist.

[0008] Die Erfassungsinformationen können Informationen über einen Zielbereich enthalten, in dem das vorbestimmte Objekt durch den zweiten Detektor nicht erfasst wird. In dieser Struktur kann die Bestimmungseinheit einen Kandidatenbereich des einen oder der mehreren Kandidatenbereiche mit einer Ähnlichkeit größer oder gleich einem vorbestimmten

Wert zu einem Zielbereich, in dem das vorbestimmte Objekt in einem vorherigen Rahmen nicht erfasst wird, ausschließen und einen oder mehrere andere Kandidatenbereiche des einen oder der mehreren Kandidatenbereiche als den Zielbereich bestimmen. Der erste Detektor kann ein erstes Erfassungskonfidenzniveau ausgeben, das eine Wahrscheinlichkeit angibt, dass das vorbestimmte Objekt in dem einen oder den mehreren Kandidatenbereichen enthalten ist. Die Bestimmungseinheit kann den Zielbereich aus einem oder mehreren Kandidatenbereichen des einen oder der mehreren Kandidatenbereiche mit einer Ähnlichkeit größer oder gleich einem vorbestimmten Wert zu einem Zielbereich, in dem das vorbestimmte Objekt in einem vorherigen Rahmen nicht erfasst wird, basierend auf einem Wert, der durch Subtrahieren eines vorbestimmten Werts von dem ersten Erfassungskonfidenzniveau erhalten wird, bestimmen. Die Bestimmungseinheit kann den Zielbereich aus einem oder mehreren anderen Kandidatenbereichen des einen oder der mehreren Kandidatenbereiche basierend auf dem ersten Erfassungskonfidenzniveau bestimmen. Diese Struktur reduziert die Anzahl von Kandidatenbereichen, die durch den zweiten Detektor verwendet werden sollen. Der zweistufige Erfassungsprozess kann somit durchgeführt werden, um die Verarbeitungszeit zu reduzieren, ohne die Erfassungsgenauigkeit zu verringern.

[0009] Der vorbestimmte Wert, der von dem ersten Erfassungskonfidenzniveau subtrahiert werden soll, kann basierend auf einer Anzahl von aufeinanderfolgenden Rahmen, in denen das vorbestimmte Objekt durch den zweiten Detektor nicht erfasst wird, bestimmt werden. Zum Beispiel kann der vorbestimmte Wert größer sein, wenn die Anzahl von aufeinanderfolgenden Rahmen zunimmt, oder der vorbestimmte Wert kann vom ersten Erfassungskonfidenzniveau zum ersten Mal subtrahiert werden, wenn die Anzahl von aufeinanderfolgenden Rahmen einen vorbestimmten Wert erreicht. Der vorbestimmte Wert, der von dem ersten Erfassungskonfidenzniveau subtrahiert werden soll, kann fest sein.

[0010] Der erste Detektor kann ein erstes Erfassungskonfidenzniveau ausgeben, das eine Wahrscheinlichkeit angibt, dass das vorbestimmte Objekt in dem einen oder den mehreren Kandidatenbereichen enthalten ist. Die Erfassungsinformationen können ein zweites Erfassungskonfidenzniveau enthalten, das durch den zweiten Detektor bestimmt wird und eine Wahrscheinlichkeit angibt, dass das vorbestimmte Objekt in dem einen oder den mehreren Kandidatenbereichen enthalten ist. Die Bestimmungseinheit kann den Zielbereich aus einem oder mehreren Kandidatenbereichen des einen oder der mehreren Kandidatenbereiche mit einer Ähnlichkeit größer oder gleich einem vorbestimmten Wert zu

einem Zielbereich, der durch die Erfassungsinformationen angegeben wird, basierend auf einem Wert, der durch Subtrahieren eines Werts, der dem zweiten Erfassungskonfidenzniveau entspricht, von dem ersten Erfassungskonfidenzniveau erhalten wird, bestimmen. Die Bestimmungseinheit kann den Zielbereich aus einem oder mehreren anderen Kandidatenbereichen des einen oder der mehreren Kandidatenbereiche basierend auf dem ersten Erfassungskonfidenzniveau bestimmen. Zum Beispiel kann der vorbestimmte Wert, der von dem ersten Erfassungskonfidenzniveau subtrahiert werden soll, größer sein, wenn das zweite Erfassungskonfidenzniveau zunimmt.

[0011] Die Erfassungsinformationen können eine Position und/oder eine Größe des Zielbereichs enthalten. Die Bestimmungseinheit kann die Ähnlichkeit basierend auf einer Position und/oder einer Größe jedes des einen oder der mehreren Kandidatenbereiche und der Position und/oder der Größe des Zielbereichs berechnen. Ein Objekt in einem Eingabebild kann wiederholt fehlerhaft erfasst werden. Diese Struktur kann jedoch eine fehlerhafte wiederholte Erfassung eines Objekts an der gleichen Position und mit der gleichen Größe effektiv reduzieren. Dies reduziert die Anzahl von Kandidatenbereichen, die durch den zweiten Detektor verwendet werden sollen. Der zweistufige Erfassungsprozess kann somit durchgeführt werden, um die Verarbeitungszeit zu reduzieren, ohne die Erfassungsgenauigkeit zu verringern.

[0012] Die Erfassungsinformationen können ein Bild enthalten, das dem Zielbereich entspricht. Die Bestimmungseinheit kann die Ähnlichkeit basierend auf dem Bild, das in den Erfassungsinformationen enthalten ist, und Bildern, die dem einen oder den mehreren Kandidatenbereichen entsprechen, berechnen. Für einen Bereich, der fehlerhaften Erfassungsinformationen entspricht, und einen Kandidatenbereich, der in der Position und Größe identisch oder ähnlich zueinander sind, kann diese Struktur eine Objekterfassung mit hoher Genauigkeit durchführen, wenn Bilder, die den zwei Bereichen entsprechen, sich weitgehend voneinander unterscheiden.

[0013] Ein Objekterfassungsverfahren gemäß einem zweiten Aspekt der vorliegenden Erfindung ist ein Verfahren zum Erfassen eines vorbestimmten Objekts aus einem Bild. Das Verfahren enthält ein Erfassen eines oder mehrerer Kandidatenbereiche einschließlich des vorbestimmten Objekts aus dem Bild, ein Bestimmen eines Zielbereichs aus dem einen oder den mehreren erfassten Kandidatenbereichen, ein Erfassen des vorbestimmten Objekts in dem Zielbereich mit einem Erfassungsalgorithmus, der sich von einem beim Erfassen des einen oder der mehreren Kandidatenbereiche verwendeten

Algorithmus unterscheidet, und ein Speichern von Erfassungsinformationen, die ein Erfassungsergebnis beim Erfassen des vorbestimmten Objekts in dem Zielbereich angeben. Das Bestimmen des Zielbereichs enthält ein Bestimmen des Zielbereichs aus dem einen oder den mehreren Kandidatenbereichen basierend auf den Erfassungsinformationen über einen Rahmen, der mindestens ein Rahmen ist, der einem aktuellen Rahmen vorausgeht.

[0014] Ein oder mehrere Aspekte der vorliegenden Erfindung können auf eine Objekterfassungseinrichtung, die mindestens eines der obigen Elemente enthält, oder auf eine Vorrichtung zum Identifizieren oder Verfolgen eines Erfassungszielobjekts, eine Bildverarbeitungsvorrichtung oder ein Überwachungssystem gerichtet sein. Ein oder mehrere Aspekte der vorliegenden Erfindung können auf ein Objekterfassungsverfahren, das mindestens einen der obigen Prozesse enthält, oder auf ein Objektividentifizierungsverfahren, ein Objektverfolgungsverfahren, ein Bildverarbeitungsverfahren oder ein Überwachungsverfahren gerichtet sein. Ein oder mehrere Aspekte der vorliegenden Erfindung können auch auf ein Programm zum Implementieren eines dieser Verfahren oder als ein nichtflüchtiges Speichermedium, das das Programm speichert, gerichtet sein. Die obigen Elemente und Prozesse können auf jede mögliche Weise miteinander kombiniert werden, um einen oder mehrere Aspekte der vorliegenden Erfindung zu bilden.

VORTEILHAFTE EFFEKTE

[0015] Die Struktur gemäß den obigen Aspekten der vorliegenden Erfindung ermöglicht eine Objekterfassung mit hoher Geschwindigkeit und mit hoher Genauigkeit.

KURZBESCHREIBUNG DER ZEICHNUNGEN

[Fig. 1] Fig. 1 ist ein Diagramm, das eine beispielhafte Verwendung einer Objekterfassung zeigt.

[Fig. 2] Fig. 2 ist ein Diagramm, das die Struktur einer Objekterfassungseinrichtung zeigt.

[Fig. 3] Fig. 3 ist ein Ablaufdiagramm einer Objekterfassung.

[Fig. 4] Fig. 4 ist ein Ablaufdiagramm einer Bestimmung.

[Fig. 5] Fig. 5 ist ein Ablaufdiagramm einer Bestimmung.

AUSFÜHRLICHE BESCHREIBUNG

(Beispielhafte Verwendung)

[0016] Unter Bezugnahme auf Fig. 1 wird eine beispielhafte Verwendung einer Objekterfassungseinrichtung gemäß einer oder mehreren Ausführungsformen der vorliegenden Erfindung beschrieben. Die Objekterfassungseinrichtung erfasst ein Zielobjekt (z. B. einen Menschen) aus einem Bild, das mit einer stationären Kamera erhalten wird, die über einem Erfassungszielbereich (z. B. einer Decke) befestigt ist. Die Objekterfassungseinrichtung umfasst zwei separate Detektoren, die ein vorhergehender Detektor und ein nachfolgender Detektor sind. Ein Objekt 101 und ein Objekt 102 sind Erfassungsziele (z. B. Menschen), die sich in einem Abbildungsbereich einer stationären Kamera 1 bewegen können. Ein Objekt 103 (z. B. eine Blume) befindet sich im Abbildungsbereich der stationären Kamera 1. Der vorhergehende Detektor in der Objekterfassungseinrichtung erfasst Kandidatenbereiche 111 bis 113 einschließlich Erfassungszielen aus einem Eingabebild. Die Kandidatenbereiche 111 bis 113 entsprechen den Objekten 101 bis 103. Das Objekt 103 ist kein zu erfassender Mensch. Wenn jedoch das Objekt 103 Merkmale aufweist, die einem Menschen ähnlich sind, wird der Kandidatenbereich 113 erzeugt. Der nachfolgende Detektor in der Objekterfassungseinrichtung führt dann eine Objekterfassung durch und zeichnet ein Erfassungsergebnis in einem Speicher auf. Der nachfolgende Detektor führt im Wesentlichen eine Objekterfassung an Zielbereichen 121 bis 123 durch, die den Kandidatenbereichen 111 bis 113 entsprechen. Wenn der vorhergehende Detektor das Objekt 103 (Blume) fehlerhaft als ein Zielobjekt erfasst, kann der nachfolgende Detektor bestimmen, dass das Objekt 103 kein Zielobjekt ist. In diesem Fall kann der vorhergehende Detektor möglicherweise das Objekt 103 weiterhin fehlerhaft erfassen. Wenn alle Kandidatenbereiche in der in Fig. 1 gezeigten Situation als Zielbereiche verwendet werden, führt der nachfolgende Detektor einen verschwenderischen Erfassungsprozess in jedem Rahmen für einen Kandidatenbereich ohne Zielobjekt durch.

[0017] In der beispielhaften Verwendung wird ein Bereich (Zielbereich), für den der nachfolgende Detektor eine Objekterfassung durchführt, aus Bereichen (Kandidatenbereichen) bestimmt, in denen der vorhergehende Detektor ein Objekt erfasst hat, basierend auf Erfassungsinformationen über einen Rahmen, der mindestens ein Rahmen ist, der dem aktuellen Rahmen vorausgeht. Wenn beispielsweise ein Kandidatenbereich im aktuellen Rahmen einem Bereich sehr ähnlich ist, in dem der nachfolgende Detektor kein Objekt in einem Rahmen erfasst hat, der mindestens ein Rahmen ist, der dem aktuellen Rahmen vorausgeht, kann der Kandidatenbereich

von Zielbereichen ausgeschlossen werden. Zielbereiche können auch aus Kandidatenbereichen basierend auf einer Erfassungsbewertung (Konfidenzniveau) des vorhergehenden Detektors bestimmt werden. Genauer gesagt, wenn der nachfolgende Detektor kein Zielobjekt in einem Bereich in einem Rahmen erfasst hat, der mindestens ein Rahmen ist, der dem aktuellen Rahmen vorausgeht, wird ein vorbestimmter Wert von der Erfassungsbewertung für den Bereich subtrahiert. Der Bereich kann dann als ein Zielbereich basierend auf der resultierenden Bewertung bestimmt werden. Der Wert, der subtrahiert werden soll, kann basierend auf der Anzahl von aufeinanderfolgenden Rahmen, in denen kein Zielobjekt erfasst wurde, festgelegt oder angepasst werden. Wie oben beschrieben, wird ein Bereich, in dem der vorhergehende Detektor ein Zielobjekt erfasst hat, selektiv von Verarbeitungszielen des nachfolgenden Detektors ausgeschlossen, wenn der Bereich einem Bereich ähnlich ist, in dem der nachfolgende Detektor kein Zielobjekt erfasst hat. Dies kann die Zeit reduzieren, die durch die Verarbeitung benötigt wird, ohne die Genauigkeit der Objekterfassung zu verringern.

(erste Ausführungsform)

<Struktur>

[0018] Fig. 2 ist ein Funktionsblockdiagramm einer Objekterfassungseinrichtung 10 gemäß der vorliegenden Ausführungsform. Die Objekterfassungseinrichtung 10 ist eine Informationsverarbeitungseinheit (Computer), die zum Beispiel eine arithmetische Einheit (eine zentrale Verarbeitungseinheit oder eine CPU oder einen Prozessor), einen Speicher, eine Speichereinheit (Speicher 16) und eine Eingabe-Ausgabe-Einheit umfasst. Die Objekterfassungseinrichtung 10 führt ein in der Speichereinheit gespeichertes Programm aus, um die Funktionen einer Bildeingabeeinheit 11, eines ersten Detektors 12, einer Bestimmungseinheit 13, eines zweiten Detektors 14, einer Ausgabereinheit 15 und anderer Einheiten zu implementieren. Diese Funktionen können teilweise oder vollständig unter Verwendung dedizierter Logikschaltungen, wie etwa einer anwendungsspezifischen integrierten Schaltung (ASIC) und eines feldprogrammierbaren Gate-Arrays (FPGA), implementiert werden.

[0019] Die Bildeingabeeinheit 11 empfängt Bilddaten von einer Kamera 20. Die aufgenommenen Bilddaten werden an den ersten Detektor 12 übertragen. Die Bilddaten können in dem Speicher 16 gespeichert werden. Obwohl Bilddaten in der vorliegenden Ausführungsform direkt von der Kamera 20 empfangen werden, können Bilddaten durch eine Vorrichtung, wie etwa eine Kommunikationsvorrichtung, oder durch ein Aufzeichnungsmedium empfangen werden. Es können beliebige Bilder eingegeben wer-

den, einschließlich RGB-Bilder, Graustufenbilder oder Bilder, die zum Beispiel Entfernung oder Temperatur darstellen.

[0020] Der erste Detektor 12 erfasst Kandidatenbereiche (Bereiche, in denen Erfassungszielobjekte wahrscheinlich enthalten sind) aus einem Eingabebild. In der vorliegenden Ausführungsform enthält der erste Detektor 12 Detektoren, die eine Haar-ähnliche Merkmalsmenge und AdaBoost zur Kandidatenbereichserfassung verwenden. Das Erfassungsergebnis wird an die Bestimmungseinheit 13 übertragen. Das Erfassungsergebnis enthält erfasste Kandidatenbereiche und kann auch die Wahrscheinlichkeit (ein erstes Erfassungskonfidenzniveau oder eine Erfassungsbewertung) von Erfassungszielobjekten enthalten, die in den Kandidatenbereichen enthalten sind. Eine solche Erfassung kann mit beliebigen Merkmalsmengen und mit Detektoren unter Verwendung eines beliebigen Lernalgorithmus durchgeführt werden. Zum Beispiel können Merkmalsmengen ein Histogramm des Gradienten (HoG), eine skaleninvariante Merkmalstransformation (SIFT), beschleunigte robuste Merkmale (SURF) und eine spärliche Merkmalsmenge enthalten. Lernalgorithmen können andere Verstärkungen als AdaBoost, eine Unterstützungsvektormaschine (SVM), ein neuronales Netzwerk und ein Entscheidungsbaumlernen enthalten.

[0021] Die Bestimmungseinheit 13 bestimmt Bereiche (Zielbereiche), für die der zweite Detektor 14 eine Erfassung durchführt, aus den durch den ersten Detektor 12 erfassten Kandidatenbereichen. In der vorliegenden Ausführungsform verwendet die Bestimmungseinheit 13 Erfassungsinformationen für den vorherigen Rahmen, die in dem Speicher 16 gespeichert sind, um Zielbereiche aus den Kandidatenbereichen zu bestimmen. Die Erfassungsinformationen enthalten Informationen über die Zielbereiche (später beschriebener fehlerhafter Erfassungsbereich), in denen der zweite Detektor 14 kein Objekt in einem Rahmen erfasst hat, der mindestens ein Rahmen ist, der dem aktuellen Rahmen vorausgeht. Die Bestimmungseinheit 13 schließt die Kandidatenbereiche mit einer Ähnlichkeit größer oder gleich einem vorbestimmten Wert zu dem fehlerhaften Erfassungsbereich aus und bestimmt die verbleibenden Bereiche als Zielbereiche. Die Bestimmungseinheit 13 gibt dann das Bestimmungsergebnis an den nachfolgenden zweiten Detektor 14 aus. Wenn das durch den ersten Detektor 12 erhaltene Erfassungsergebnis das obige erste Erfassungskonfidenzniveau enthält, kann die Bestimmungseinheit 13 die Kandidatenbereiche ähnlich dem fehlerhaften Erfassungsbereich aus den Kandidatenbereichen mit dem ersten Erfassungskonfidenzniveau größer oder gleich dem vorbestimmten Wert ausschließen und bestimmt die verbleibenden Bereiche als Zielbereiche.

[0022] Der zweite Detektor 14 führt eine Objekterfassung für die durch die Bestimmungseinheit 13 bestimmten Zielbereiche durch. Das Erfassungsergebnis enthält Informationen, die angeben, ob Erfassungszielobjekte in den Zielbereichen enthalten sind, und kann auch beispielsweise die Wahrscheinlichkeit (ein zweites Erfassungskonfidenzniveau oder eine Erfassungsbewertung) von Erfassungszielobjekten enthalten, die in den Zielbereichen enthalten sind. In der vorliegenden Ausführungsform zeichnet der zweite Detektor 14 nach der Objekterfassung beide oder entweder die Position oder die Größe des Zielbereichs, von dem bestimmt wurde, dass er kein Erfassungszielobjekt enthält, als Erfassungsinformationen in den Speicher 16 auf. In einigen Ausführungsformen kann der zweite Detektor 14 die Erfassungsinformationen (beide oder entweder die Position oder die Größe) über alle durch die Bestimmungseinheit 13 bestimmten Zielbereiche in den Speicher 16 aufzeichnen. In der vorliegenden Ausführungsform enthält der zweite Detektor 14 Detektoren, die ein tiefes Lernen zur Objekterfassung verwenden. Die Detektoren können beliebige bekannte Techniken des tiefen Lernens verwenden, wie beispielsweise ein neuronales Faltungsnetzwerk (CNN), ein rekurrentes neuronales Netzwerk (RNN), einen gestapelten Autocodierer (SAE) und ein tiefes Glaubensnetzwerk (DBN). Der zweite Detektor 14 kann kein Detektor sein, der ein tiefes Lernen verwendet. Der Erfassungsalgorithmus mit dem zweiten Detektor 14 kann eine genauere Erfassung mit einer größeren Rechenmenge als der Erfassungsalgorithmus mit dem ersten Detektor 12 ermöglichen.

[0023] Die Ausgabeinheit 15 gibt ein Erfassungsergebnis über ein durch den zweiten Detektor 14 erfasstes Objekt aus. Beispielsweise gibt die Ausgabeinheit 15 Ergebnisinformationen aus, die angeben, dass ein Objekt in dem Kandidatenbereich erfasst wurde, wobei das Konfidenzniveau des Erfassungsergebnisses durch den zweiten Detektor 14 größer oder gleich einem Schwellenwert ist. Die Ergebnisinformationen können die Informationen über die Kandidatenbereiche mit dem Konfidenzniveau kleiner als der Schwellenwert ausschließen. Alle Informationen können als Erfassungsergebnisinformationen verwendet werden. Beispiele für die Erfassungsergebnisinformationen für die Gesichtserfassung beinhalten Informationen über einen Gesichtsbereich, ein Konfidenzniveau, eine Gesichtsausrichtung, ein Alter, ein Geschlecht, eine menschliche Rasse und einen Gesichtsausdruck.

<Prozess>

[0024] Fig. 3 ist ein Ablaufdiagramm des Gesamtprozesses einer Objekterfassung, der durch die Objekterfassungseinrichtung 10 durchgeführt wird. Die Objekterfassungseinrichtung 10 wird nun

gemäß dem Ablaufdiagramm in Fig. 3 ausführlich beschrieben.

«S31: Bildeingabe»

[0025] In Schritt S31 ermittelt die Objekterfassungseinrichtung 10 ein Bild (Eingabebild). Das Eingabebild kann von der Kamera 20 durch die Bildeingabeinheit 11, von einem anderen Computer durch eine Kommunikationseinheit 104 oder von dem Speicher 16 ermittelt werden.

«S32: erste Erfassung»

[0026] In Schritt S32 erfasst der erste Detektor 12 Kandidatenbereiche (Bereiche, in denen erwartet wird, dass ein Erfassungszielobjekt enthalten ist) aus dem Eingabebild (erste Erfassung). In der vorliegenden Ausführungsform verwendet der erste Detektor 12 die Haar-ähnliche Merkmalsmenge als eine Bildmerkmalsmenge und AdaBoost als einen Lernalgorithmus. Das Erfassungsergebnis bei der ersten Erfassung kann die Wahrscheinlichkeit (ein erstes Erfassungskonfidenzniveau oder eine Erfassungsbewertung) von Erfassungszielobjekten enthalten, die in den Kandidatenbereichen enthalten sind, sowie die obigen Kandidatenbereiche.

«S33: Bestimmung»

[0027] In Schritt S33 schließt die Bestimmungseinheit 13 die in Schritt S32 erfassten Kandidatenbereiche mit einer Ähnlichkeit größer oder gleich einem vorbestimmten Wert zu einem fehlerhaften Erfassungsbereich aus und bestimmt die verbleibenden Kandidatenbereiche als Zielbereiche. Der fehlerhafte Erfassungsbereich ist ein Bereich, in dem kein Objekt bei der zweiten Erfassung in einem Rahmen erfasst wurde, der mindestens ein Rahmen ist, der dem aktuellen Rahmen vorausgeht. Die zweite Erfassung wird später beschrieben. Die Bestimmungseinheit 13 schließt die in Schritt S32 erfassten Kandidatenbereiche ähnlich dem fehlerhaften Erfassungsbereich aus und gibt die verbleibenden Kandidatenbereiche als Zielbereiche aus.

[0028] Die in Schritt S33 durchgeführte Bestimmung wird nun unter Bezugnahme auf Fig. 4 ausführlich beschrieben. Fig. 4 ist ein Ablaufdiagramm, das die Bestimmung in der vorliegenden Ausführungsform zeigt. Die Bestimmungseinheit 13 erhält zunächst Erfassungsinformationen (die Position und die Größe eines fehlerhaften Erfassungsbereichs) aus dem Speicher 16 (S41). Die Bestimmungseinheit 13 kann fehlerhafte Erfassungsinformationen über einen unmittelbar vorhergehenden Rahmen allein oder mehrere vorhergehende Rahmen ermitteln. Die Bestimmungseinheit 13 berechnet dann eine Ähnlichkeit zu dem fehlerhaften Erfassungsbereich für jeden des einen oder der mehreren Kandidaten-

bereiche (S42). In der vorliegenden Ausführungsform wird Intersection over Union (IoU) als ein Index für eine Ähnlichkeit zwischen den Bereichen verwendet. IoU ist ein Wert, der durch Teilen eines Schnittbereichs von zwei Bereichen durch den Verbindungsbereich der zwei Bereiche berechnet wird. IoU kann ein beliebiger Wert zwischen 0 und 1 sein. IoU ist 1 für zwei Bereiche, die sich vollständig überlappen, und 0 für zwei Bereiche, die sich überhaupt nicht überlappen. Die Positionen und die Größen eines Kandidatenbereichs und eines fehlerhaften Erfassungsbereichs können verwendet werden, um IoU zu berechnen. Die Bestimmungseinheit 13 bestimmt, ob IoU größer oder gleich einem vorbestimmten Schwellenwert T1 ist (S43). Die Bestimmungseinheit 13 schließt dann die Kandidatenbereiche mit IoU größer oder gleich dem vorbestimmten Schwellenwert T1 aus und gibt die verbleibenden Bereiche als Zielbereiche aus (S44).

«S34 bis S36: zweite Erfassung»

[0029] In Schritt S34 bestimmt der zweite Detektor 14, ob ein Erfassungszielobjekt in einem oder mehreren in Schritt S33 ausgegebenen Zielbereichen enthalten ist (zweite Erfassung). In der vorliegenden Ausführungsform verwendet der zweite Detektor 14 Diskriminatoren, die durch ein mehrschichtiges neuronales Netzwerk, das CNN genannt wird, zur Objekterfassung gelernt haben.

[0030] In Schritt S35 bestimmt der zweite Detektor 14, ob einer der Zielbereiche in Schritt S34 als kein Erfassungszielobjekt enthaltend bestimmt wurde.

[0031] In Schritt S36 zeichnet der zweite Detektor 14 Informationen über den Zielbereich, von dem bestimmt wurde, dass er kein Erfassungszielobjekt enthält, als Erfassungsinformationen in den Speicher 16 auf. In der vorliegenden Ausführungsform werden die Position und die Größe des Zielbereichs, von dem bestimmt wurde, dass er kein Erfassungszielobjekt enthält, als die Erfassungsinformationen in den Speicher 16 aufgezeichnet.

«S37: Erfassungsergebnisausgabe»

[0032] In Schritt S37 gibt die Ausgabereinheit 15 ein Erfassungsergebnis über den Bereich aus, in dem in Schritt S34 ein Objekt erfasst wurde. Die Ausgabereinheit 15 gibt Ergebnisinformationen aus, die angeben, dass ein Erfassungszielobjekt in einem Erfassungszielbereich erfasst wurde, wobei das Konfidenzniveau (zweites Erfassungskonfidenzniveau) des Erfassungsergebnisses durch den zweiten Detektor 14 größer oder gleich einem Schwellenwert ist. Die Ergebnisinformationen können die Informationen über die Erfassungszielbereiche mit dem Konfidenzniveau kleiner als der Schwellenwert abschließen.

<Vorteilhafte Wirkungen der ersten Ausführungsform>

[0033] Ein Objekt in einem Eingabebild kann wiederholt fehlerhaft erfasst werden. Die Struktur gemäß der vorliegenden Ausführungsform kann jedoch eine fehlerhafte Erfassung eines Objekts an der gleichen Position und mit der gleichen Größe effektiv reduzieren. Dies reduziert die Anzahl von Kandidatenbereichen (Zielbereichen), die durch den zweiten Detektor verwendet werden sollen. Der zweistufige Erfassungsprozess kann somit durchgeführt werden, um die Verarbeitungszeit zu reduzieren, ohne die Erfassungsgenauigkeit zu verringern.

(zweite Ausführungsform)

[0034] In der oben beschriebenen ersten Ausführungsform wird die Ähnlichkeit in Schritt S33 basierend auf den Positionen und den Größen der Kandidatenbereiche und der fehlerhaften Erfassungsbereiche bestimmt. In der vorliegenden Ausführungsform wird die Ähnlichkeit in Schritt S33 durch Musterabgleich zwischen einem Bild, das einem Kandidatenbereich entspricht, und einem Bild, das einem fehlerhaften Erfassungsbereich entspricht, bestimmt. Die Verarbeitung, die die gleiche wie in der oben beschriebenen ersten Ausführungsform ist, wird nicht beschrieben, und eine Bestimmung (S33), die sich von der in der ersten Ausführungsform beschriebenen Verarbeitung unterscheidet, wird beschrieben.

<Bestimmung (S33)>

[0035] Fig. 5 ist ein Flussdiagramm einer Bestimmung in Schritt S33 in der vorliegenden Ausführungsform. Die Bestimmungseinheit 13 ermittelt zunächst Erfassungsinformationen aus dem Speicher 16 (S51). In der vorliegenden Ausführungsform enthalten die Erfassungsinformationen ein Bild, das einem fehlerhaften Erfassungsbereich entspricht. Die Bestimmungseinheit 13 führt einen Musterabgleich an jedem Bild, das einem oder mehreren Kandidatenbereichen entspricht, unter Verwendung eines Bilds, das einem fehlerhaften Erfassungsbereich entspricht, durch (S52). Die Bestimmungseinheit 13 bestimmt, ob die Ähnlichkeit zwischen den durch den Musterabgleich ermittelten Bildern größer oder gleich einem Schwellenwert T2 ist (S53), und schließt dann die Kandidatenbereiche mit der Ähnlichkeit größer oder gleich dem vorbestimmten Schwellenwert T2 aus und gibt die verbleibenden Bereiche als Zielbereiche aus (S54).

<Vorteilhafte Wirkungen der zweiten Ausführungsform>

[0036] Für einen fehlerhaften Erfassungsbereich und einen Kandidatenbereich, die in der Position

und Größe identisch oder ähnlich zueinander sind, kann diese Struktur eine Objekterfassung mit hoher Genauigkeit durchführen, wenn Bilder, die den zwei Bereichen entsprechen, sich weitgehend voneinander unterscheiden. Wenn beispielsweise ein Erfassungszielobjekt das in **Fig. 1** gezeigte Objekt 103 überlappt, wird die Ähnlichkeit basierend auf Bildern berechnet, und somit kann ein Bereich, der der Position des Erfassungszielobjekts entspricht, erfolgreich als ein Zielbereich erfasst werden.

(Modifikationen)

[0037] Obwohl die Bestimmungseinheit 13 die Kandidatenbereiche ähnlich dem fehlerhaften Erfassungsbereich aus den Kandidatenbereichen ausschließt und die verbleibenden Bereiche als Zielbereiche in der oben beschriebenen ersten und zweiten Ausführungsform bestimmt, kann die Bestimmungseinheit 13 einen Zielbereich mit einem beliebigen anderen Verfahren bestimmen. Wenn beispielsweise der erste Detektor 12 das oben beschriebene erste Erfassungskonfidenzniveau ausgibt, bestimmt die Bestimmungseinheit 13 einen Kandidatenbereich mit dem ersten Erfassungskonfidenzniveau größer oder gleich einem vorbestimmten Schwellenwert T3 als einen Zielbereich. Die Bestimmungseinheit 13 kann einen Kandidatenbereich als einen Zielbereich mit einer Ähnlichkeit größer oder gleich einem vorbestimmten Schwellenwert T4 zu einem fehlerhaften Erfassungsbereich bestimmen, wenn ein solcher Kandidatenbereich einen Wert aufweist, der durch Subtrahieren eines vorbestimmten Werts von dem ersten Erfassungskonfidenzniveau größer oder gleich dem vorbestimmten Schwellenwert T3 erhalten wird.

[0038] Der vorbestimmte Wert, der von dem ersten Erfassungskonfidenzniveau subtrahiert werden soll, kann mit einem beliebigen Verfahren bestimmt werden. Ein Wert, der von dem Konfidenzniveau subtrahiert wird, kann fest sein. In einigen Ausführungsformen kann der vorbestimmte Wert, der von dem Konfidenzniveau subtrahiert wird, basierend auf der Anzahl von aufeinanderfolgenden Rahmen, in denen der zweite Detektor 14 kein Zielobjekt erfasst hat, bestimmt werden. Zum Beispiel kann der vorbestimmte Wert größer sein, wenn die Anzahl von aufeinanderfolgenden Rahmen zunimmt, oder der vorbestimmte Wert kann von dem ersten Erfassungskonfidenzniveau zum ersten Mal subtrahiert werden, wenn die Anzahl von aufeinanderfolgenden Rahmen einen vorbestimmten Wert erreicht. Ferner kann, wenn der zweite Detektor 14 das zweite Erfassungskonfidenzniveau ausgibt, der vorbestimmte Wert, der von dem Konfidenzniveau subtrahiert wird, basierend auf dem zweiten Erfassungskonfidenzniveau bestimmt werden. Zum Beispiel bestimmt die Bestimmungseinheit 13 einen Kandidatenbereich mit dem ersten Konfidenzniveau größer

oder gleich dem Schwellenwert T3 als einen Zielbereich. Die Bestimmungseinheit 13 kann einen Kandidatenbereich als einen Zielbereich mit einer Ähnlichkeit größer oder gleich dem vorbestimmten Schwellenwert T4 zu einem fehlerhaften Erfassungsbereich bestimmen, wenn ein solcher Kandidatenbereich einen Wert aufweist, der durch Subtrahieren eines Werts basierend auf dem zweiten Erfassungskonfidenzniveau größer oder gleich dem vorbestimmten Schwellenwert T3 erhalten wird. Zum Beispiel kann der vorbestimmte Wert, der von der Konfidenz subtrahiert wird, größer sein, wenn die zweite Erfassungskonfidenz zunimmt.

[0039] Obwohl IoU als ein Index einer Ähnlichkeit zwischen Bereichen in der oben beschriebenen ersten Ausführungsform verwendet wird, können beliebige andere Indizes verwendet werden. Solche Indizes einer Ähnlichkeit enthalten das Verhältnis oder die Differenz zwischen den Größen von Bereichen, die Differenz zwischen den Positionen von Bereichen (z. B. den Koordinaten der Zentren) oder eine beliebige Kombination davon.

[0040] Obwohl ein Musterabgleich verwendet wird, um eine Ähnlichkeit zwischen Bildern in der oben beschriebenen zweiten Ausführungsform zu bestimmen, können beliebige andere Techniken verwendet werden. Zum Beispiel können Differenzen in Farbinformationen oder in Luminanzinformationen in Bildern als ein Index einer Ähnlichkeit verwendet werden.

BEZUGSZEICHENLISTE

10	Objekterfassungseinrichtung
11	Bildeingabeeinheit
12	erster Detektor
13	Bestimmungseinheit
14	zweiter Detektor
15	Ausgabeeinheit
16	Speicher
1, 20	Kamera
101, 102, 103	Objekt
111, 112, 113	Kandidatenbereich
121, 122, 123	Zielbereich

ZITATE ENTHALTEN IN DER BESCHREIBUNG

Diese Liste der vom Anmelder aufgeführten Dokumente wurde automatisiert erzeugt und ist ausschließlich zur besseren Information des Lesers aufgenommen. Die Liste ist nicht Bestandteil der deutschen Patent- bzw. Gebrauchsmusteranmeldung. Das DPMA übernimmt keinerlei Haftung für etwaige Fehler oder Auslassungen.

Zitierte Patentliteratur

- JP 2006293720 [0002]
- JP 2019021001 [0002]

Patentansprüche

1. Objekterfassungseinrichtung zum Erfassen eines vorbestimmten Objekts aus einem Bild, wobei die Objekterfassungseinrichtung umfasst:
 einen ersten Detektor, der konfiguriert ist, um einen oder mehrere Kandidatenbereiche einschließlich des vorbestimmten Objekts aus dem Bild zu erfassen;
 eine Bestimmungseinheit, die konfiguriert ist, um einen Zielbereich aus dem einen oder den mehreren durch den ersten Detektor erfassten Kandidatenbereichen zu bestimmen;
 einen zweiten Detektor, der konfiguriert ist, um das vorbestimmte Objekt in dem Zielbereich mit einem Erfassungsalgorithmus zu erfassen, der sich von einem durch den ersten Detektor verwendeten Erfassungsalgorithmus unterscheidet; und
 einen Speicher, der konfiguriert ist, um Erfassungsinformationen zu speichern, die ein durch den zweiten Detektor für den Zielbereich ermitteltes Erfassungsergebnis angeben,
 wobei die Bestimmungseinheit den Zielbereich aus dem einen oder den mehreren Kandidatenbereichen basierend auf den Erfassungsinformationen über einen Rahmen bestimmt, der mindestens ein Rahmen ist, der einem aktuellen Rahmen vorausgeht.

2. Objekterfassungseinrichtung nach Anspruch 1, wobei die Erfassungsinformationen Informationen über einen Zielbereich enthalten, in dem das vorbestimmte Objekt durch den zweiten Detektor nicht erfasst wird.

3. Objekterfassungseinrichtung nach Anspruch 2, wobei die Bestimmungseinheit einen Kandidatenbereich des einen oder der mehreren Kandidatenbereiche mit einer Ähnlichkeit größer oder gleich einem vorbestimmten Wert zu einem Zielbereich, in dem das vorbestimmte Objekt in einem vorherigen Rahmen nicht erfasst wird, ausschließt und einen oder mehrere andere Kandidatenbereiche des einen oder der mehreren Kandidatenbereiche als den Zielbereich bestimmt.

4. Objekterfassungseinrichtung nach Anspruch 2, wobei
 der erste Detektor ein erstes Erfassungskonfidenzniveau ausgibt, das eine Wahrscheinlichkeit angibt, dass das vorbestimmte Objekts in dem einen oder den mehreren Kandidatenbereichen enthalten ist, und
 die Bestimmungseinheit den Zielbereich aus einem oder mehreren Kandidatenbereichen des einen oder der mehreren Kandidatenbereiche mit einer Ähnlichkeit größer oder gleich einem vorbestimmten Wert zu einem Zielbereich, in dem das vorbestimmte Objekt in einem vorherigen Rahmen nicht erfasst wird, basierend auf einem Wert, der durch Subtrahieren eines vorbestimmten Werts von dem ersten

Erfassungskonfidenzniveau erhalten wird, bestimmt und den Zielbereich aus einem oder mehreren anderen Kandidatenbereichen des einen oder der mehreren Kandidatenbereiche basierend auf dem ersten Erfassungskonfidenzniveau bestimmt.

5. Objekterfassungseinrichtung nach Anspruch 4, wobei der vorbestimmte Wert basierend auf einer Anzahl von aufeinanderfolgenden Rahmen, in denen das vorbestimmte Objekt durch den zweiten Detektor nicht erfasst wird, bestimmt wird.

6. Objekterfassungseinrichtung nach Anspruch 4, wobei der vorbestimmte Wert fest ist.

7. Objekterfassungseinrichtung nach Anspruch 1, wobei
 der erste Detektor ein erstes Erfassungskonfidenzniveau ausgibt, das eine Wahrscheinlichkeit angibt, dass das vorbestimmte Objekt in dem einen oder den mehreren Kandidatenbereichen enthalten ist, die Erfassungsinformationen ein zweites Erfassungskonfidenzniveau enthalten, das durch den zweiten Detektor bestimmt wird und eine Wahrscheinlichkeit angibt, dass das vorbestimmte Objekt in dem einen oder den mehreren Kandidatenbereichen enthalten ist, und
 die Bestimmungseinheit den Zielbereich aus einem oder mehreren Kandidatenbereichen des einen oder der mehreren Kandidatenbereiche mit einer Ähnlichkeit zu einem Zielbereich, die durch die Erfassungsinformationen angegeben wird, größer oder gleich einem vorbestimmten Wert basierend auf einem Wert, der durch Subtrahieren eines Werts, der dem zweiten Erfassungskonfidenzniveau entspricht, von dem ersten Erfassungskonfidenzniveau erhalten wird, bestimmt und den Zielbereich aus einem oder mehreren anderen Kandidatenbereichen des einen oder der mehreren Kandidatenbereiche basierend auf dem ersten Erfassungskonfidenzniveau bestimmt.

8. Objekterfassungseinrichtung nach einem der Ansprüche 3 bis 7, wobei
 die Erfassungsinformationen eine Position und/oder eine Größe des Zielbereichs enthalten und
 die Bestimmungseinheit die Ähnlichkeit basierend auf einer Position und/oder einer Größe jedes des einen oder der mehreren Kandidatenbereiche und der Position und/oder der Größe des Zielbereichs berechnet.

9. Objekterfassungseinrichtung nach einem der Ansprüche 3 bis 7, wobei
 die Erfassungsinformationen ein Bild enthalten, das dem Zielbereich entspricht,
 und
 die Bestimmungseinheit die Ähnlichkeit basierend auf dem Bild, das in den Erfassungsinformationen enthalten ist, und Bildern, die dem einen oder den

mehreren Kandidatenbereichen entsprechen, berechnet.

10. Objekterfassungsverfahren zum Erfassen eines vorbestimmten Objekts aus einem Bild, wobei das Verfahren umfasst:

Erfassen eines oder mehrerer Kandidatenbereiche einschließlich des vorbestimmten Objekts aus dem Bild;

Bestimmen eines Zielbereichs aus dem einen oder den mehreren erfassten Kandidatenbereichen;

Erfassen des vorbestimmten Objekts in dem Zielbereich mit einem Erfassungsalgorithmus, der sich von einem beim Erfassen des einen oder der mehreren Kandidatenbereiche verwendeten Algorithmus unterscheidet; und

Speichern von Erfassungsinformationen, die ein Erfassungsergebnis beim Erfassen des vorbestimmten Objekts in dem Zielbereich angeben,

wobei das Bestimmen des Zielbereichs ein Bestimmen des Zielbereichs aus dem einen oder den mehreren Kandidatenbereichen basierend auf den Erfassungsinformationen über einen Rahmen, der mindestens ein Rahmen ist, der einem aktuellen Rahmen vorausgeht, umfasst.

11. Programm zum Bewirken, dass ein Computer Operationen durchführt, die in dem Objekterfassungsverfahren nach Anspruch 10 enthalten sind.

Es folgen 5 Seiten Zeichnungen

Anhängende Zeichnungen

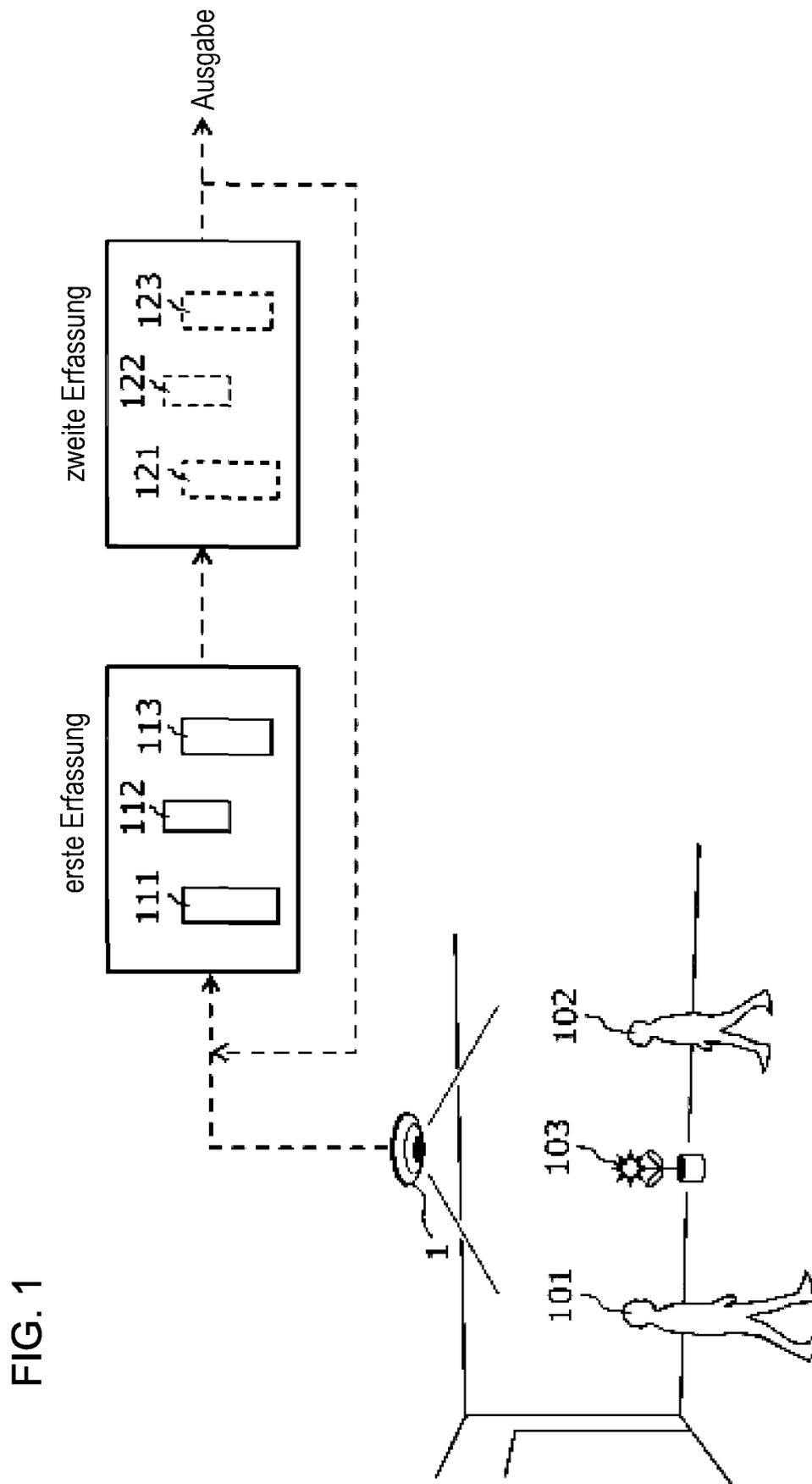


FIG. 2

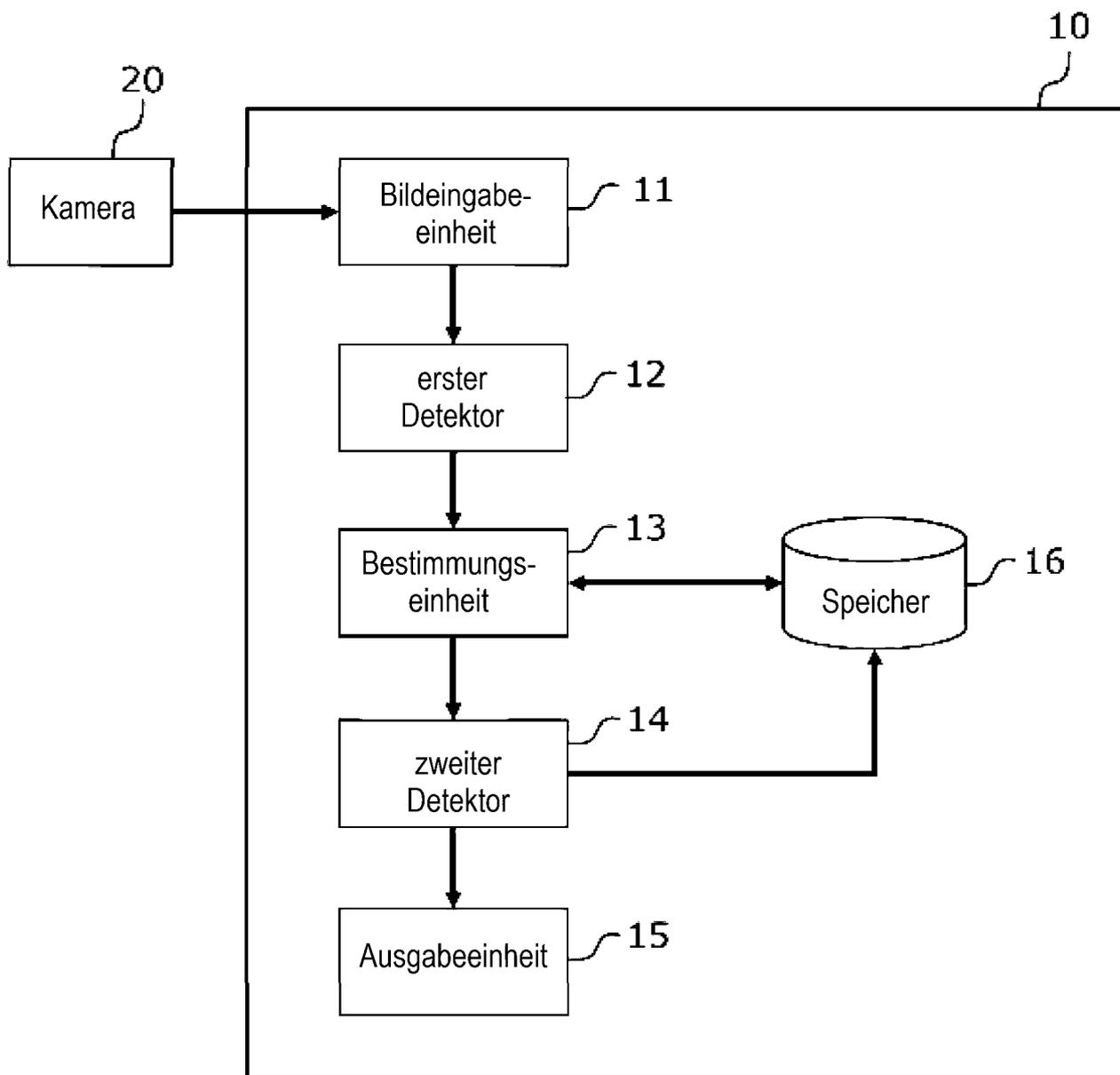


FIG. 3

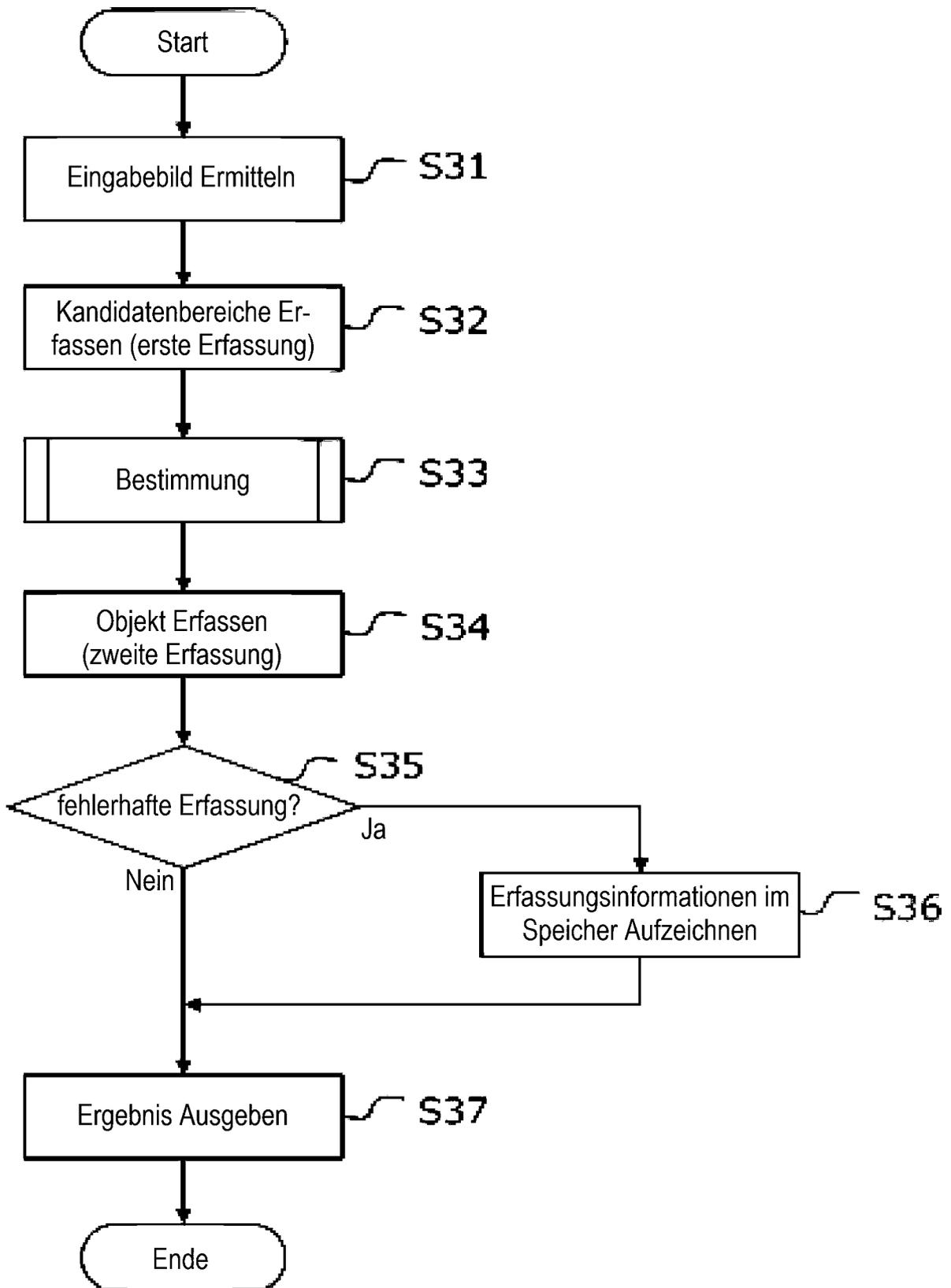


FIG. 4

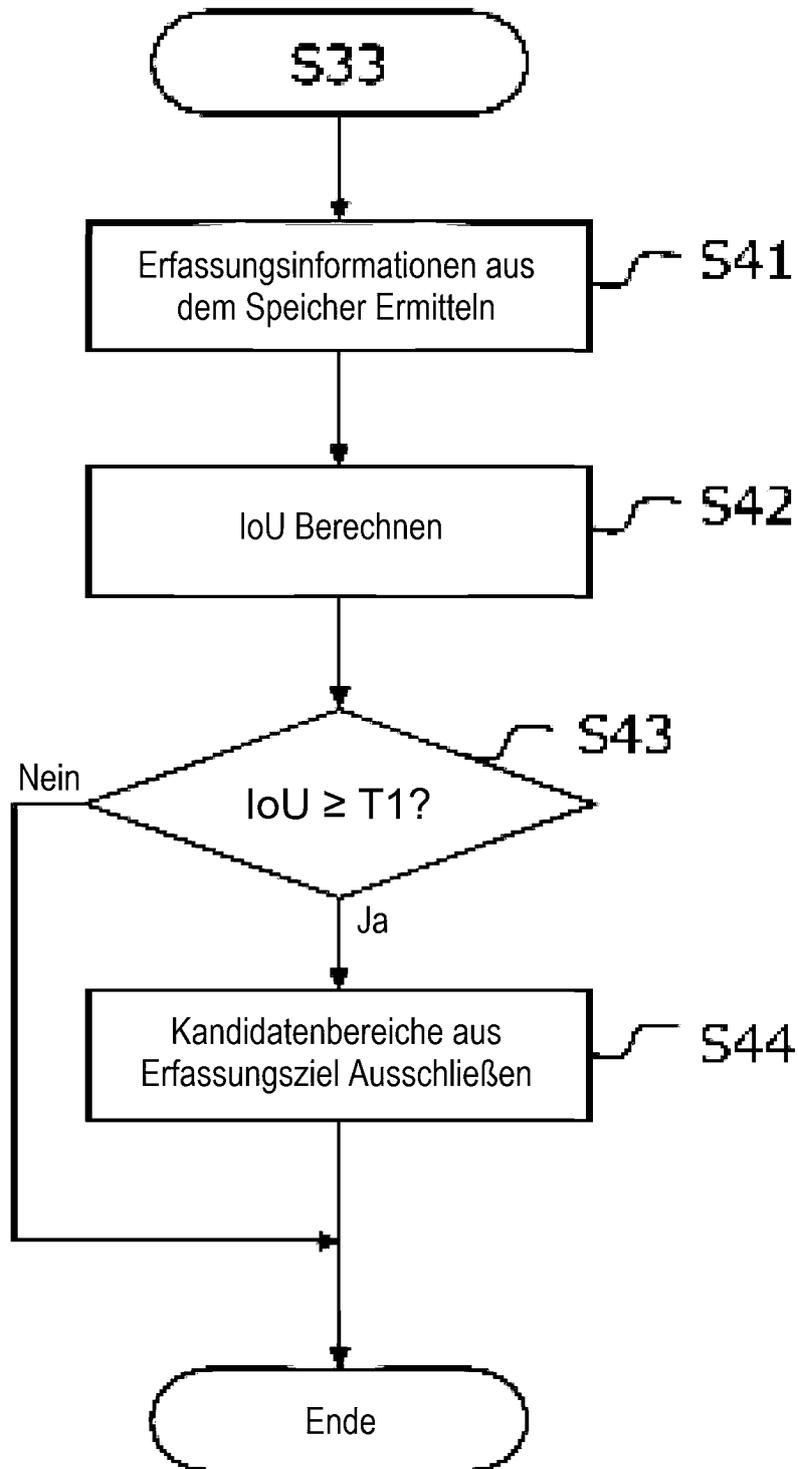


FIG. 5

